Почему аргумент Бострома некорректен: проблема отсутствующего субъекта

Деконструкция симуляционной трилеммы через анализ её скрытой эпистемологической посылки

Трилемма Ника Бострома, сформулированная в две тысячи третьем году, остаётся одним из самых обсуждаемых философских аргументов нашего времени. Напомню её суть. Одно из трёх утверждений должно быть истинным.

Либо цивилизации, как правило, гибнут раньше, чем обретают способность создавать полноценные симуляции сознания. Либо те, что выживают, по каким-то причинам отказываются это делать. Либо мы с высокой вероятностью уже находимся внутри такой симуляции.

Двадцать лет философы, физики, специалисты по искусственному интеллекту спорят о том, какой вариант вероятнее. Написаны сотни работ. Предложены вероятностные модели, контраргументы, уточнения. Но при всём разнообразии позиций есть одна вещь, которую, насколько мне известно, никто не поставил под сомнение: сама корректность вопроса.

Я хочу показать, что вопрос поставлен некорректно. Не потому, что логика Бострома ошибочна — она безупречна. А потому, что в основании аргумента лежит посылка, которую никто не проверял.

Кто решает?

Посмотрите на формулировки внимательно. Цивилизация вымирает. Цивилизация отказывается. Цивилизация создаёт симуляции. Во всех трёх случаях действующее лицо одно и то же — цивилизация. Она выступает как субъект, как агент, принимающий решения или терпящий последствия.

Но что такое цивилизация как субъект? Это не метафора и не условность. Для того чтобы трилемма работала, цивилизация должна быть способна на нечто вроде коллективного выбора. Не голосования, не рыночного равновесия, а именно выбора — осознанного действия в отношении собственной природы.

Здесь возникает вопрос, который я считаю ключевым и который, по-видимому, до сих пор не был задан в явной форме: существует ли такой субъект?

Пять шагов к ответу

Я предлагаю следующую цепочку рассуждений. Каждый шаг в ней, как мне представляется, может быть проверен независимо.

Первое. Аргумент Бострома требует цивилизационного субъекта. Без агента, способного вымирать, отказываться или создавать, три варианта трилеммы теряют смысл. Это не интерпретация — это прямое следствие структуры аргумента.

Второе. Если цивилизация является субъектом, она должна обладать чем-то вроде органа самопознания. Субъект, который не способен к рефлексии о собственной природе, не может «решать», создавать ему симуляции или нет. Это было бы всё равно что приписывать способность к выбору системе, у которой нет механизма осознания альтернатив.

Третье. Такого органа не существует. Я утверждаю это как верифицируемый факт и приглашаю любого оппонента его опровергнуть. На сегодняшний день нет ни одного института, ни одной структуры, ни одной междисциплинарной программы, которая была бы уполномочена исследовать вопрос о природе цивилизации как целого — не в рамках философии, не в рамках физики, не в рамках информатики, а именно как целостную проблему, требующую одновременного учёта всех этих перспектив.

Вопрос Бострома не принадлежит ни одной дисциплине. Он попадает в пространство между ними — туда, где нет ни кафедр, ни журналов, ни экспертов. Я называю это пространство Центральным Коридором Распознавания. Это не теория и не метафора. Это наблюдение: существуют вопросы, которые видны из пространства между дисциплинами, но для которых внутри дисциплинарной системы нет адресата. Вопрос Бострома — один из них.

Четвёртое. Если органа самопознания нет, то цивилизация не является субъектом. Она является средой. Совокупность индивидов, каждый из которых мыслит в рамках своей дисциплины, своего опыта, своей нейрологической ограниченности, не образует агента. Нейрология здесь однозначна: человеческий мозг организован вокруг индивидуального восприятия. Коллективное сознание — красивая метафора, но у неё нет нейрологического субстрата. Мы не муравьи и не клетки единого организма. Мы — отдельные люди в отдельных комнатах.

Пятое. Если субъекта нет, аргумент Бострома теряет адресата. Трилемма описывает выбор агента, которого не существует. Она не ложна и не истинна. Она некорректна — в том же смысле, в каком некорректен вопрос «какого цвета зависть?». Грамматически безупречен, логически пуст.

Что можно проверить

Центральное утверждение этой работы — отсутствие института, способного целостно исследовать вопрос о природе цивилизации — поддаётся прямой проверке. Достаточно попытаться найти такой институт.

Ни один университет мира не имеет кафедры, чьей задачей было бы исследование вопроса «является ли наша цивилизация симуляцией» как целостной проблемы, а не как темы для семинара по философии сознания. Ни один государственный орган не

уполномочен ставить этот вопрос. Ни одна грантовая программа не финансирует его исследование в целостной постановке.

Это не упущение и не недоработка. Я полагаю, что это структурное свойство существующей системы знания. Система организована так, что целостные вопросы неизбежно дробятся на дисциплинарные фрагменты. Каждый фрагмент исследуется добросовестно, но целое не собирается нигде и никем. Коридор между комнатами пуст.

К чему это ведёт

Я не утверждаю, что мы не живём в симуляции. Я не утверждаю обратного. Я утверждаю нечто более скромное и, как мне кажется, более важное: у нас нет субъекта, от имени которого этот вопрос мог бы быть поставлен.

Двадцать лет дискуссии вокруг трилеммы Бострома были, по существу, дискуссией внутри некорректно поставленного вопроса. Каждый, кто выбирал один из трёх вариантов, неявно принимал существование цивилизационного субъекта — и тем самым воспроизводил ошибку, заложенную в основании аргумента.

Это наблюдение, в свою очередь, ставит собственный вопрос: почему за двадцать лет никто не заметил отсутствующего субъекта? Возможно, ответ в том, что для этого нужно стоять не внутри какой-либо дисциплины, а в пространстве между ними. В коридоре, из которого видны все комнаты, но который сам — ничей.

Работа выполнена с использованием метода ЦКР, то есть Центрального Коридора Распознавания — подхода, исследующего не содержание дисциплинарных ответов, а условия, при которых определённые вопросы могут или не могут быть поставлены.